JOURNAL OF COMPUTER AND SYSTEM SCIENCES 7, 199–217 (1973)

On the Extension of Gladkij's Theorem and the Hierarchies of Languages

YOSHIHIDE IGARASHI AND NAMIO HONDA

Research Institute of Electrical Communication, Tohoku University, Sendai, Japan 980

Received August 2, 1972

Let T be a computable, monotonic increasing function from non-negative integers to positive integers. Then it is said that $\alpha \subseteq \Sigma^*$ is in a class L_T if there exists a phrase structure grammar G, which generates all words of length n in α within length $T(n)$ of derivations for each n. A main result of this paper is an extension of the A. V. Gladkij's Nonlinear Theorem on Context-Sensitive Grammars. Our extended theorem is as follows: Let φ be a function defined on the set of all strings on an alphabet $\Sigma = \{a_1, a_2\}$, taking as values non-void subsets of Σ^*. Let $\mathscr{L}_\varphi$ be a language $\{xb\varphi(x)\, bx \mid x \in \Sigma^*\}$, where b is a symbol not in Σ. Let $\mathscr{F}_\varphi$ be a function from Σ^* to positive integres defined by $\mathscr{F}_\varphi(x) = \min\{\mid xb\varphi(x)\, bx \mid\}$, and let f_φ be a function from non-negative integers to positive integers defined by $f_\varphi(n) = \max\{\mathscr{F}_\varphi(x) \mid x \in \Sigma^*, \mid x \mid = n\}$. If T is a time function such that $\lim_{n\to\infty}(T(f_\varphi(n))/n^2) = 0$, then L_T does not contain $\mathscr{L}_\varphi$. From this result, an open problem proposed by R. V. Book are solved. Moreover from this result, it is shown that there exist infinitely long chains of distinct complexity classes between certain two distinct complexity classes.

INTRODUCTION

As the study of formal languages and automata has been enlarged, many ideas have been proposed to define complexity classes of formal languages. In particular, many authors defined the complexity classes by Turing machines used as recognizers. Complexity classes of formal languages based upon the derivational complexities were also defined. A. V. Gladkij derived the interesting theorem on the derivational complexity classes, called the Nonlinear Theorem of Context-Sensitive Grammars, in 1964. In this paper, we extend the Gladkij's Nonlinear Theorem to a more general form. As a corollary of our theorem, an open problem proposed by R. V. Book are solved. Some hierarchy problems are also solved by using our extended theorem.

This paper is divided into four sections. In Section 1, we review the basic concepts of phrase structure languages, and define derivational complexity measures. In Section 2, we describe the concept given by A. V. Gladkij in slightly general form. In Section 3, we derive the extended theorem of Gladkij's Nonlinear Theorem of Context-Sensitive Grammars. From this theorem the open problem of R. V. Book, some nonclosure proporties and some relations between classes of languages are solved. In Section 4, the existence of infinitly long chains of distinct complexity classes between two certain classes are described.

1. Preliminary Definitions

In this section we recall some basic concepts about grammars and languages for the necessity of an understanding of this paper, and give derivational complexity measures. Let S^* be a set of all finite sequences of elements from a set S including the empty word ϵ, and let $S^+ = S^* - \{\epsilon\}$. As usual, if u and v are strings, then uv denotes the concatenation of u and v (u followed by v), and $|u|$ is the length of u. If $u = u_1u_2u_3$, u_2 is called a subword of u. If $u = u_1u_2u_3$ and $|u| \gneq |u_2|$, u_2 is called a proper subword of u. If $u = u_1u_2$, u_1 is called a prefix of u and u_2 is called a suffix of u.

Definition 1. A phrase structure grammar (abbreviated PSG) is a 4-tuple $G = (V, \Sigma, P, \sigma)$, where

(1) V is a finite set.

(2) $\Sigma \subseteq V$ is an alphabet.

(3) P is a finite set of the form $u \to v$, with u in $(V - \Sigma)^+$ and v in V^*.

(4) σ is in $V - \Sigma$.

Elements of $V - \Sigma$ are called variables, and elements of Σ are called terminals. Elements $u \to v$ of P are called production rules. σ is called a initial symbol.

Definition 2. Let $G = (V, \Sigma, P, \sigma)$ be a PSG. For w and y in V^*, write $w \Rightarrow_G y$ (or $w \Rightarrow y$ when G is understood) if there exist z_1, z_2, u and v such that $w = z_1uz_2$, $y = z_1vz_2$ and $u \to v$ is in P. For w and y in V^*, write $w \overset{*}{\Rightarrow}_G y$ (or $w \overset{*}{\Rightarrow} y$ when G is understood) if either $w = y$ or there exist $w_0, w_1, \ldots, w_r$ such that $w_0 = w$, $w_r = y$, and $w_i \Rightarrow w_{i+1}$ for each i. The sequence $w_0, \ldots, w_r$ is called a derivation and is denoted by $w_0 \Rightarrow w_1 \Rightarrow \cdots \Rightarrow w_r$. The subset of Σ^*, $L(G) = \{x \text{ in } \Sigma^* \mid \sigma \overset{*}{\Rightarrow} x\}$ is called a phrase structure language (abbreviated PSL). If $L(G) = L(G')$, G is equivalent to G'.

Definition 3. Let $G = (V, \Sigma, P, \sigma)$ be a PSG. The length of a derivation $\xi \Rightarrow w_1 \Rightarrow \cdots \Rightarrow w_n$ in G is the number of rewriting steps and is denoted by $\mathrm{dl}(\xi \Rightarrow w_1 \Rightarrow \cdots \Rightarrow w_n)$. The length of a minimum length derivation of w from ξ in G is denoted by $\mathrm{dlm}(\xi \overset{*}{\Rightarrow} w)$.

Definition 4. A context-sensitive grammar (abbreviated CSG) is a PSG $G = (V, \Sigma, P, \sigma)$ in which each production rule is of the form $u \to v$, with $|u| \leqslant |v|$.

A set $\alpha \subseteq \Sigma^*$ is called a context-sensitive language (abbreviated CSL) if there exists a CSG such that $\alpha = L(G)$.

DEFINITION 5. A context-sensitive with erasing grammar (abbreviated CSEG) is a PSG $G = (V, \Sigma, P, \sigma)$ in which each production rule is of the form $uAv \to u\eta v$, where A is in $V - \Sigma$, η is in V^*, u and v are in $(V - \Sigma)^*$.

For any PSG G, there exists an equivalent CSEG G' such that for each derivation $u \overset{*}{\Rightarrow} v$ of G $\text{dlm}(u \overset{*}{\Rightarrow}_G v) \geqslant (\text{dlm}(u \overset{*}{\Rightarrow}_{G'} v))/k$, where k is an appropriate positive integer. Hence it is clear that the class of PSL's coincides with the class of languages generated by CSEG's.

DEFINITION 6. A context-free grammar (abbreviated CFG) is a PSG $G = (V, \Sigma, P, \sigma)$ in which each production rule is of the form $u \to v$, with u in $V - \Sigma$ and v in V^*. A set $\alpha \subseteq \Sigma^*$ is called a context-free language (abbreviated CFL) if there exists a CFG G such that $\alpha = L(G)$.

DEFINITION 7. $T(n)$ is a time function if and only if $T(n)$ is a computable function from nonnegative integers into positive integers such that $T(n) \leqslant T(n + 1)$ for each nonnegative integer n.

DEFINITION 8. Let T be a time function. Then we say that a PSG $G = (V, \Sigma, P, \sigma)$ is a T-derivable grammar if and only if $T(|w|) \geqslant \text{dlm}(\sigma \overset{*}{\Rightarrow} w)$ for each w in $L(G)$. A set $\alpha \subseteq \Sigma^*$ is said to be a T-derivable language if and only if α is generated by some T-derivable grammar. The class of all T-derivable languages is denoted by L_T.

In this paper, unless stated otherwise, $T(n) = f(n)$ means $T(n) = [f(n)]$, where $[r]$ is the smallest positive integer m such that $m \geqslant r$.

The complexities of derivations by A. V. Gladkij [11] and R. V. Book [2] can be defined in our notations as the following two definitions.

DEFINITION 9. Let T be a time function and let $G = (V, \Sigma, P, \sigma)$ be a PSG. Then we say that a PSG G is a T-bounded grammar if and only if $T(|w|) \geqslant \text{dlm}(\sigma \overset{*}{\Rightarrow} w)$ for each w in V^* derived from σ. A set $\alpha \subseteq \Sigma^*$ is said to be a T-bounded language if and only if α is generated by some T-bounded grammar. The class of all T-bounded languages is denoted by $L_{\langle T \rangle}$.

DEFINITION 10. Let T be a time function and let $G = (V, \Sigma, P, \sigma)$ be a CSG. Then we say that G is a T-derivable CSG if and only if $T(|w|) \geqslant \text{dlm}(\sigma \overset{*}{\Rightarrow} w)$ for each w in $L(G)$, and we say that G is a T-bounded CSG if and only if $T(|w|) \geqslant \text{dlm}(\sigma \overset{*}{\Rightarrow} w)$ for each w in V^* derived from σ. A set $\alpha \subseteq \Sigma^*$ is said to be a T-derivable CSL if and only if α is generated by some T-derivable CSG. A set $\alpha \subseteq \Sigma^*$ is said to be a T-bounded CSL if and only if α is generated by some T-bounded CSG. The class of all T-derivable CSL's is denoted by L_{-T-}, and the class of all T-bounded CSL's is denoted by $L_{\langle\langle T\rangle\rangle}$.

From the above definitions, it is clear that for any time function $T, L_T \supseteq L_{\langle T\rangle} \supseteq L_{\langle\langle T\rangle\rangle}$, $L_T \supseteq L_{-T-} \supseteq L_{\langle\langle T\rangle\rangle}$. It is known that there exists a time function T satisfying $L_T \supsetneqq L_{\langle T\rangle}$ [14]. We shall give an example of language α and a time function T such that α is in $L_T - L_{\langle T\rangle}$.

EXAMPLE. Let $G = (V, \Sigma, P, \sigma_0)$ be a PSG, where $V = \{\sigma_0, \sigma_1, \hat{a}, \hat{b}, a_{\#} a_e, \bar{b}, b_{\#}, a, b, d\}$, $\Sigma = \{d\}$ and P is a set of the following production rules:

$$
\begin{array}{ll}
\sigma_0 \to \hat{b}\hat{a} & b\hat{a} \to a_{\#} a_e\, d \\
\sigma_0 \to \sigma_1 \hat{a} & b a_{\#} \to a_{\#} a d \\
\sigma_1 \to \sigma_1 b & b a_e \to a a_e\, d \\
\sigma_1 \to \hat{b} b & ba \to aab \\
\hat{b}\hat{a} \to a_{\#} a \bar{b} & \hat{b} a_{\#} \to a_{\#} a \hat{b} \\
\hat{b} a \to a a \hat{b} & a b_{\#} \to b_{\#} b_{\#} a \\
\hat{b} a_e \to a a \bar{b} & a_{\#} b_{\#} \to d d a_{\#} \\
a \bar{b} \to b_{\#} \bar{b} d & a_{\#} \bar{b} \to ddd
\end{array}
$$

Then $L(G) = \{d^{f(f(t))+f(t)+t-1} \mid t \geqslant 1\}$, where $f(t) = 2^t$. Let $T(n) = \max\{1, \{\mathrm{dlm}(\sigma_0 \overset{*}{\Rightarrow} w) \mid w \in L(G), \mid w \mid \leqslant n\}\}$. Then $L_T \ni L(G)$, but $L_{\langle T\rangle} \not\ni L(G)$.

The proof of the above example is given in [14]. For any time function T satisfying the condition $\lim_{n\to\infty}(T(n)/n) = 0$, $L_T = L_{\langle T\rangle}$, since both L_T and $L_{\langle T\rangle}$ are the classes of all finite set of terminal words, respectively. It is still open whether for all time function T satisfying the condition $\lim_{n\to\infty}(T(n)/n) \gneqq 0$, $L_{\langle T\rangle}$ is properly contained in L_T. The time function T in the above example satisfies neither the condition $\lim_{n\to\infty}(T(n)/n) = 0$ nor the condition $\lim_{n\to\infty}(T(n)/n) \gneqq 0$.

The following two definitions are also given in [2] and [11].

DEFINITION 11. Let G be a PSG and let $W = (\theta_0 \Rightarrow \theta_1 \Rightarrow \cdots \Rightarrow \theta_n)$ be a derivation in G. By production sequence for the derivation W we mean a sequence of pairs

$$\{(\theta_{i-1} = B_i \xi_i C_i, \theta_i = B_i \eta_i C_i)\}_{i=1}^{n},$$

where $\xi_i \to \eta_i$ is a production rule of G. For each i three cases are possible:

(1) $\mid B_{i+1}\xi_{i+1} \mid \leqslant \mid B_i \mid$,

(2) $\mid B_{i+1} \mid \geqslant \mid B_i \eta_i \mid$,

(3) $\mid B_{i+1}\xi_{i+1} \mid \gneqq \mid B_i \mid$ and $\mid B_{i+1} \mid \lneqq \mid B_i \eta_i \mid$.

In case (3), it is said that the $(i+1)$th step is connected with the ith step.

DEFINITION 12. Let $G = (V, \Sigma, P, \sigma)$ be a PSG. A derivation from σ is called a proper derivation. A derivation in G is called a connected derivation if it has at least one production sequence such that each step is connected with the previous step. G is a connected grammar if each proper derivation in G is connected.

2. AUXILIARY CONCEPTS AND LEMMAS

In this section, we prepare some concepts and lemmas for the necessity of an understanding of the following sections. These concepts were introduced by A. V. Gladkij [11]. Since minor changes are made in the definitions, the lemmas in this section are not the same as those of A. V. Gladkij. However the proof is essentially the same. Most of these lemmas are proved in [13].

LEMMA 1. *Let* $G = (V, \Sigma, P, \sigma)$ *be a* PSG. *For any constant* $k \gneq 0$, *we can effectively find a connected grammar* $G' = (V', \Sigma, P', \sigma')$ *such that for all* $w \in V^*$, $[k \cdot \mathrm{dl}(\sigma \overset{*}{\Rightarrow}_G w)] \geqslant \mathrm{dlm}(\sigma' \overset{*}{\Rightarrow}_{G'} w)$. [13]

LEMMA 2. *Let* α *be a* T*-derivable language. Then for any positive computable number* k, α *is a* kT*-derivable language.* [2][13]

Each production rule of a PSG can be replaced by a derivation in which some finite number of production rules satisfying the conditions of CSEG and connected grammar are applied. Hence the next lemma follows from lemma 1.

LEMMA 3. *Let* $G = (V, \Sigma, P, \sigma)$ *be a* PSG. *Then there exist a connected* CSEG G' *and positive integer* k *such that for each derivation* $\xi \overset{*}{\Rightarrow} \eta$ *in* G,

$$\mathrm{dlm}(\xi \underset{G'}{\overset{*}{\Rightarrow}} \eta) \leqslant k\mathrm{dlm}(\xi \underset{G}{\overset{*}{\Rightarrow}} \eta) \qquad \text{and} \qquad L(G) = L(G').$$

From Lemma 2 and Lemma 3 we may consider that every PSL is generated by the CSEG in the following argument. From the proof of Lemma 1 (given in [13]) and the definition of CSEG's, the following proposition is clear: If G is a connected T-derivable CSEG, we can construct a connected and rT-derivable CSEG G' equivalent to G such that there exists exactly one production sequence for each proper derivation in G', where r is an appropriate positive integer. Hence without loss of generality, we may suppose that each CSEG G is connected and has exactly one production sequence for each derivation in G.

DEFINITION 13. Suppose that $G = (V, \Sigma, P, \sigma)$ is a CSEG, $W = (\theta_0 \Rightarrow \theta_1 \Rightarrow \cdots \Rightarrow \theta_n)$ is a derivation in G, and $\{(\theta_{i-1} = u_i\xi_iv_i\,,\ \theta_i = u_i\eta_iv_i)\}_{i=1}^n$ is a production sequence for W. Furthermore, suppose that for each $i(1 \leqslant i \leqslant n)$, $\xi_i = \rho_i A_i \psi_i$ and $\eta_i = \rho_i \nu_i \psi_i$, where $A_i \in V - \Sigma$. Then we say that at the ith step the symbol A_i in θ_{i-1} is replaced by ν_i and the remaining symbols in θ_{i-1} are copied. Let $\theta_j = \tau w \zeta$ and let $\theta_{j+r} = \tau' w' \zeta' (1 \leqslant j \leqslant n,\ 1 \leqslant r \leqslant n - j)$. We inductively define the relation, "w is the (kth) exact ancestor of w'" or in other words, "w' is the (kth) exact descendant of w" as follows:

(i) For $k = 1$, the above relation holds if $w = b_1 \cdots b_s$ and $w' = B_1 \cdots B_s$, where each $b_m (1 \leqslant m \leqslant s)$ is a symbol in V and each B_m is the result of replacing or copying b_m.

(ii) Suppose that the relation has been defined for $k = r - 1$. For $k = r$, the relation holds if $\theta_{j+r-1} = \tau'' w'' \zeta''$, and w'' is the exact $(r - 1)$th descendant of w and the exact 1st ancestor of w'.

If the kth exact descendant of w is contained in w' as a proper subword of w' and if the concatenation of w, and the last symbol of τ or the first symbol of ζ gives a string whose kth exact descendant is not a contained in w' as a subword of w', then w is called the (kth) inexact ancestor of w', or in other words, w' is the (kth) inexact descendant of w. Furthermore, we say that $w = \epsilon$ is the (kth) inexact ancestor of w', or in other words, w' is the (kth) inexact descendant of $w = \epsilon$ if one of the following two conditions is satisfied: (i') w' is contained in the kth exact descendant of D as a proper subword, where D is in V. (ii') w' is contained in the kth exact descendant of DC as a proper subword, and neither the kth exact descendant of D nor the kth exact descendant of C is contained in w' as a subword, where D and C are in V. An exact ancestor or an inexact ancestor is called an ancestor, and an exact descendant or an inexact descendant is called a descendant.

Definition 14. Let $G = (V, \Sigma, P, \sigma)$ be a CSEG and let $g = \max\{|\xi|, |\eta|, |\xi \to \eta \in P\}$. Suppose that $W = (\theta_0 \Rightarrow \cdots \Rightarrow \theta_n)$ is a derivation in G which is a part of a proper derivation in G. Suppose that $\theta_0 = \theta_0' \theta_0''$ and $|\theta_0'| = e$. Suppose also θ_j' and θ_j'' are the jth exact descendants of θ_0' and θ_0'', respectively. Let the interface between θ_0' and θ_0'' and also between θ_j' and θ_j'' be denoted by the number e. Suppose that $\bar{\theta}_j'$ and $\bar{\theta}_j''$ are respectively the longest suffix of θ_j' and the longest prefix of θ_j'' whose lengths do not exceed g. Then $\bar{\theta}_j' \bar{\theta}_j''$ is called the jth zone of influence of the point e. If we want to indicate the distance of each symbol in θ_j from the point e, we write $\bar{\theta}_j' = a_{-q} \cdots a_{-1}$ and $\bar{\theta}_j'' = a_1 \cdots a_r$, where each $a_m (-q \leqslant m \leqslant r,\ m \neq 0)$ is in V and m indicates the distance of the symbol a_m from the point e.

Next we shall define the trace of derivation. ξ and η in a production rule $\xi \to \eta$ are called a left-hand side and a right-hand side of the rule, respectively.

Definition 15. Let G be a CSEG and let each production rule of G be indicated by an appropriate natural number. Let W be a derivation in G which is a part of a proper derivation in G. Suppose that $\{j_1, \dots, j_s\}$ is a set consisting of step numbers of the derivation W in which the left-hand side of the applied rules are entirely contained in the corresponding (j_tth) zones of influence of the point e. The trace of the derivation W at the point e is a sequence $\{(k_1, m_1), \dots, (k_s, m_s)\}$, where k_t is distance between the point e and the symbol replaced at the j_tth step, and m_t is the index of the rule applied at this j_tth step. The number of pairs k and m in the trace is called the length of the trace.

A. V. Gladkij proved a proposition on CSG's which is called the Replacement Lemma.

LEMMA. *The notion in the Replacement Lemma is similar to that of "crossing sequence" of one-tape off-line Turing machine computations introduced by F. C. Hennie.* [12] *The proposition on* CSEG's *corresponding to the Replacement Lemma on* CSG's *is also true.*

LEMMA 4. *Let G be a connected* CSEG. *Suppose that $W = (\theta_0 \Rightarrow \theta_1 \Rightarrow \cdots \Rightarrow \theta_n)$ and $W' = (\eta_0 \Rightarrow \eta_1 \Rightarrow \cdots \Rightarrow \eta_m)$ are derivations in G which are parts of proper derivations in G, respectively. Suppose $\theta_0 = \theta_0'\theta_0''$, $\theta_n = \theta_n'\theta_n''$, $\eta_0 = \eta_0'\eta_n''$ and $\eta_m = \eta_m'\eta_m''$, where θ_n', θ_n'', η_m' and η_m'' are exact descendants of θ_0', θ_0'', η_0' and η_0'', respectively. If the trace of the derivation W at the interface of θ_0' and θ_0'' coincides with the trace of the derivation W' at the interface of η_0' and η_0'', then the string $\theta_n'\eta_m''$ is derived from $\theta_0'\eta_0''$ in G.*

Proof. Suppose $\{(k_1, u_1), \dots, (k_s, u_s)\}$ is the common trace, and $(j_1, \dots, j_s)$ and $(j_1', \dots, j_s')$ are the corresponding sequences of steps in the derivations W and W', respectively. Suppose $k_1 \lneq 0$ (proof is similar even when $k_1 \gneq 0$). Then at the steps of derivation W preceding the j_1th (if there are any), only the string θ_0' can be transformed. If for some $t(1 \leqslant t \leqslant s-1)$ the numbers k_t and k_{t+1} have different signs (one is positive and other is negative), then $j_{t+1} = j_t + 1$ (notice that G is connected). If k_t and k_{t+1} are both positive (negative), then between j_tth and j_{t+1}th steps only the string $\theta_j''(\theta_j')$ can be transformed. The same is true for the derivation W'. Therefore $\theta_n'\eta_m''$ can be derived from $\theta_0'\eta_0''$ as follows:

First, θ_0' is transformed as in the derivation W until the j_1th step. The transformation of the left part from the interface is continued as in the derivation W until the first step for which k_t and k_{t+1} have different signs ($k_t \lneq 0$). Then, if k_t and k_{t+1} have different signs, transform $\eta_{j_t'}'' = \eta_0''$ as in the derivation W' until the first step for which k_{t+r} is positive and k_{t+r+1} is negative ($r \geqslant 1$). Continue the transformation in this way. In this derivation, the left part and the right part from the interface are independently transformed as in the derivation W and the derivation W', respectively. Eventually we will obtain the string $\theta_n'\eta_m''$ from the string $\theta_0'\eta_0''$ in G. Q.E.D.

3. Extension of the Nonlinear Theorem

This section is the main part of this paper. A. V. Gladkij derived a very interesting theorem, called the Nonlinear Theorem on CSG's, using the Replacement Lemma on CSG's [11]. It was shown that, for example, the language $\{wcw \mid w \in \Sigma^*\}$ is not contained in $L_{\langle\langle i \rangle\rangle}$, where c is a symbol not in Σ and $i(n) = n$. In this section, we will extend the Gladkij's Nonlinear Theorem to a more general form. An open problem given by R. V. Book [2] will be solved from a corollary of our extended theorem. Some relations between classes of languages will also be solved from our extended theorem.

Definition 16. Let φ be a function defined on the set of all strings on an alphabet $\Sigma = \{a_1, a_2\}$, taking nonvoid subsets of Σ^* as values. Let $\mathscr{L}_\varphi$ be a language $\{xb\varphi(x)\, bx \mid x \in \Sigma^*\}$, where b is a symbol not in Σ. Let $\mathscr{F}_\varphi$ be a function from Σ^* to positive integers defined by $\mathscr{F}_\varphi(x) = \min\{\mid xbybx \mid y \in \varphi(x)\}$ and let f_φ be a function from nonnegative integers to positive integers defined by

$$f_\varphi(n) = \max\{\mathscr{F}_\varphi(x) \mid \mid x \mid = n, x \in \Sigma^*\}.$$

Theorem 5. *Let T be a time function such that $\lim_{n\to\infty} (T(f_\varphi(n))/n^2) = 0$. Then L_T does not contain $\mathscr{L}_\varphi$.*

Proof. Assume that there exists a T-derivable connected CSEG $G = (V, \{a_1, a_2, b\}, P, \sigma)$ such that $L(G) = \mathscr{L}_\varphi$, where T satisfies the condition

$$\lim_{n\to\infty} (T(f_\varphi(n))/n^2) = 0. \tag{1}$$

Then we will be lead to a contradiction.

The proof of this theorem is very long. A brief sketch of how the argument proceeds is as follows: We shall define a series of subset of $\mathscr{L}_\varphi$, $\mathscr{L}_\varphi \supseteq \alpha_0{}^t \supseteq \alpha_1{}^t \cdots \supseteq \alpha_i{}^t$, satisfying certain conditions, and we shall give each lower bound of the number of elements of each subset in the series. Applying Lemma 4 at appropriate points in an element of the last subset $\alpha_i{}^t$ in the series, we shall have a contradiction such that an element not in $\mathscr{L}_\varphi$ is in $\mathscr{L}_\varphi$. The complete proof is as follows.

Let k be the smallest integer such that 2^k is not smaller than the number of elements in V. We introduce the integer parameters t and n such that $n = 5\,kt$. We may assume that t is sufficiently large. Let $g(x)$ be an element of $\varphi(x)$ whose length is not larger than $f_\varphi(\mid x \mid) - 2 \mid x \mid - 2$, and let $\alpha_0{}^j$ be a subset of $\mathscr{L}_\varphi$ defined (from the definition of φ, g is a total function from Σ^* to Σ^*):

$$\alpha_0{}^j = \{xbg(x)\, bx \mid \mid x \mid = 2n\}.$$

From the definition of $\alpha_0{}^t$, $\#(\alpha_0{}^t) = 2^{2n}$, where $\#(\alpha)$ denotes the number of elements in a set α. For each string x on $\{a_1, a_2\}$ of length $2n$, we put $Y(x) = xbg(x)\, bx$, and represent $Y(x)$ in the form

$$Y(x) = Y_1(x)\, Y_c(x)\, Y_2(x), \qquad \text{where} \qquad |\, Y_1(x)| = |\, Y_2(x)| = n.$$

Let $D_{Y(x)} = ((Y^0(x) = \sigma) \Rightarrow Y^1(x) \Rightarrow \cdots \Rightarrow (Y^m(x) = Y(x)))$ be a derivation such that $m \leqslant T(f_\varphi(2n))$. For any $p_1 \gneq 0$ and a sufficiently large n, it follows from the equation (1) that

$$m \lneq p_1 n^2. \tag{2}$$

For any $j (0 \leqslant j \leqslant m)$, $Y^j(x)$ is of the form $Y^j(x) = Y_1{}^j(x)\, y_1{}^j(x)\, Y_c{}^j(x)\, y_c{}^j(x)\, Y_2{}^j(x)$, where $Y_i{}^j(x)$ $(i = 1, 2, c)$ is an ancestor of $Y_i(x)$, and $y_1{}^j(x)$ and $y_c{}^j(x)$ are in $V \cup \{\epsilon\}$, respectively. Since $Y_i{}^j$ is not in general the exact ancestor of Y^j, $y_1{}^j$ and $y_c{}^j$ are necessarily introduced. Suppose j_0 is the least number of j's for which $|\, Y_c{}^j(x)| \geqslant t$. Only the following two cases can arise. We will show that both of them lead to contradictions.

Case 1. For at least half of all strings $Y(x)$ of $\alpha_0{}^t$, the length of $Y_1{}^{j_0}(x)$ and $Y_2{}^{j_0}(x)$ are both longer than t.

Case 2. For at least half of all strings $Y(x)$ of $\alpha_0{}^t$, either the length of $Y_1{}^{j_0}(x)$ or the length of $Y_2{}^{j_0}(x)$ is less than t.

Proof of Case 1. Let $\alpha_1{}^t$ be the set of all strings of α_0 which satisfy the condition of the case 1. $\#(\alpha_1{}^t) \geqslant 2^{10kt}/2$, because $\#(\alpha_0{}^t) = 2^{10kt}$ and in the case 1, at least half of all strings in $\alpha_0{}^t$ satisfy the condition of the Case 1. $2^{10kt}/2 \geqslant 2^{39kt/4}$, because t is sufficiently large. Then $\#(\alpha_1{}^t) \geqslant 2^{39kt/4}$.

For each element $Y(x)$ of $\alpha_1{}^t$, we define $Q(x) = (Y_1^{j_0}(x))^{(t)} y_1^{j_0}(x)\; Y_c^{j_0}(x)\, y_c^{j_0}(x)^{(t)}(Y_2^{j_0}(x))$, where ${}^{(t)}w(w^{(t)})$ denotes the prefix (suffix) of length t of w. Let $\alpha_2{}^t$ be the largest subset of $\alpha_1{}^t$ such that every element $Y(x)$ of $\alpha_2{}^t$ has the same $Q(x)$. Since for each $Y(x)$ in $\alpha_1{}^t$, $3t + s + 2 \geqslant |\, Q(x)| \geqslant 3t$, the number of different $Q(x)$ is at most

$$\sum_{h=0}^{s} (2^k)^{3t+2+h}, \qquad \text{where } s = \max\{|\,\eta\,|\mid \xi \to \eta \in p\}.$$

Then, we have $\#(\alpha_2) \geqslant 2^{39kt/4}/(s + 2)(2^k)^{3t+2+s} \geqslant 2^{26kt/4}$.

Let $D_x{}'$ be a derivation which is the segment of the derivation $D_{Y(x)}$ beginning with $Y^{j_0}(x)$. We denote by ρ_e the length of the trace of the derivation $D_x{}'$ at the point e in the string $Y^{j_0}(x)$. The summation of the length of the trace of $D_x{}'$ at each point in $Y^{j_0}(x)$ is in proportion to the length of the derivation. Then

$$\sum_{e=0}^{|Y^{j_0}(x)|} \rho_e \leqslant rm, \tag{3}$$

where r is a constant depending only on the grammar G. From (2) and (3), for any $p_2 \geq 0$ and sufficiently large t, there exist points e_1, e_2, e_3 satisfying the inequality (4) in each of the segments $(Y_1^{j_0}(x))^{(t)}$, $Y_c^{j_0}(x)$ and ${}^{(t)}(Y_2^{j_0}(x))$.

$$\rho_e \sim p_2 t. \tag{4}$$

A tuple (e_1, e_2, e_3) means that e_1 is a point in $(Y_1^{j_0}(x))^{(t)}$, e_2 is a point in $Y_c^{j_0}(x)$ and e_3 is a point in ${}^{(t)}(Y_2^{j_0}(x))$.

Let $\hat{\alpha}_2^t(e_1, e_2, e_3)$ be a set of all elements of α_2^t satisfying the inequality (4) at a given points e_1, e_2 and e_3. Since ${}_{3t+3+s}C_3 \geqslant (3t)^3$, there exists a (e_1, e_2, e_3) such that $\#(\hat{\alpha}_2^t(e_1, e_2, e_3)) \geqslant 2^{26kt/4}/(3t)^3 \geqslant 2^{25kt/4}$. Let (e_1, e_2, e_3) be one of tuples satisfying the above condition, and let α_3^t be the set $\hat{\alpha}_2^t(e_1, e_2, e_3)$. Let α_4^t be the largest subset of α_3^t such that every element $Y(x)$ of α_4^t have the same traces of the derivation D_x' at the respective points e_1, e_2 and e_3. From (4), $\#(\alpha_4^t) \geqslant 2^{24kt/4}$. For an arbitrary string $Y(x)$ in α_4^t, we represent $Y^{j_0}(x)$ in the form $Y^{j_0}(x) = z_1(x)\, z_2(x)\, z_3(x)\, z_4(x)$, where the point of interface between $z_i(x)$ and $z_{i+1}(x) (i = 1, 2, 3)$ is e_i. Since α_4^t is a subset of α_2^t, all three points e_1, e_2 and e_3 lie in $Q(x)$. Therefore the segments $z_2(x)$ and $z_3(x)$ for all x respectively coincide. That is, for each pair of $Y(x_1) \in \alpha_4^t$ and $Y(x_2) \in \alpha_4^t$, $z_2(x_1) = z_2(x_2) = \hat{z}_2$ and $z_3(x_1) = z_3(x_2) = \hat{z}_3$.

The exact descendant of the segment $z_i(x)$ in $Y(x)$ is denoted by $U_i(x)$ $(i = 1, 2, 3, 4)$. Let α_5^t be the largest subset of α_4^t such that for any $Y(x_1)$ and $Y(x_2)$ in α_5^t the segments $U_i(x_1)$ and $U_i(x_2)$ have the equal length $(i = 1, 2, 3, 4)$. The number of ways of choosing lengths of $U_1(x)$, $U_2(x)$, $U_3(x)$ and $U_4(x)$ is not larger than the number of ways we can insert three marks in the string of length $f_\varphi(2n)$. The order of this number is t^3 and so is less than $2^{1kt/4}$. Therefore $\#(\alpha_5^t) \geqslant 2^{23kt/4}$. We represent the length of $U_i(x)$ by d_i, where $Y(x)$ is in α_5^t. Then we have the following inequalities:

$$d_1 \leqslant 5kt, \qquad d_4 \leqslant 5kt; \tag{5}$$

$$d_1 + d_2 \geqslant 5kt, \qquad d_3 + d_4 \geqslant 5kt. \tag{6}$$

Moreover, at least one of the two inequalities $d_1 + d_2 \geqslant 10kt$ and $d_3 + d_4 \geqslant 10kt$ holds. Without loss of generality we may assume the former inequality (the proof for the latter case is similar).

Let α_6^t be the largest subset of α_5^t such that for any $Y(x_1)$ and $Y(x_2)$ in α_6^t, $U_1(x_1) = U_1(x_2)$. Since the number of different elements of length d_1 is 2^{d_1}, we have the following inequality:

$$\#(\alpha_6^t) \geqslant 2^{23kt/4}/2^{d_1} = 2^{(23kt/4)-d_1}. \tag{7}$$

From (5) and (7), $\#(\alpha_6^t) \geqslant 2^{3kt/4}$. Let $\hat{U}_1$ denote the common string $U_1(x)$ for $Y(x)$ in α_6^t. Now we will show that there exist $Y(x_1)$ and $Y(x_2)$ in α_6^t satisfying the relation

$$u_3(x_1)\, u_4(x_1) \neq u_3(x_2)\, u_4(x_2). \tag{8}$$

Suppose that for any $Y(x_1)$ and $Y(x_2)$ in $\alpha_6{}^t$, the relation

$$u_3(x_1)\,u_4(x_1) = u_3(x_2)\,u_4(x_2) \tag{9}$$

holds. Recall that $Y(x_i) = x_i b g(x_i)\, b x_i$ and $|\,x_i\,| = 10kt$. If $d_1 + d_3 + d_4 \geqslant 10kt$, then the relation (8) and the relation $u_1(x_1) = u_1(x_2)$ lead to the relation $x_1 = x_2$. Thus $\alpha_6{}^t$ is a singleton set. If $d_1 + d_3 + d_4 \lneqq 10kt$, then $\#(\alpha_6{}^t) \leqslant 2^{10kt-d_1-d_3-d_4} \leqslant 2^{5kt-d_1}$. These results contradict the inequality (7). Therefore there exist $Y(x_1)$ and $Y(x_2)$ in $\alpha_6{}^t$ satisfying relation (8).

Let $Y(x_1)$ and $Y(x_2)$ be strings in $\alpha_6{}^t$ satisfying relation (8). We have the following equalities:

$$Y^{j_0}(x_1) = z_1(x_1)\hat{z}_2\hat{z}_3 z_4(x_1), \tag{10}$$

$$Y^{j_0}(x_2) = z_1(x_2)\hat{z}_2\hat{z}_3 z_4(x_2), \tag{11}$$

$$Y(x_1) = \hat{U}_1 U_2(x_1)\, U_3(x_1)\, U_4(x_1), \tag{12}$$

and

$$Y(x_2) = \hat{U}_1 U_2(x_2)\, U_3(x_2)\, U_4(x_2). \tag{13}$$

Applying Lemma 4 at the point e_1 in (10), (11), (12), and (13), we get (14).

$$z_1(x_1)\, \hat{z}_2\hat{z}_3 z_4(x_2) \overset{*}{\Rightarrow} \hat{U}_1 U_2(x_2)\, U_3(x_2)\, U_4(x_2). \tag{14}$$

Applying Lemma 4 at the point e_2 in (10), (12), and (14), we get (15):

$$z_1(x_1)\, \hat{z}_2\hat{z}_3 z_4(x_1) \overset{*}{\Rightarrow} \hat{U}_1 U_2(x_2)\, U_3(x_1)\, U_4(x_1) \tag{15}$$

Therefore $\hat{U}_1 U_2(x_2)\, U_3(x_1)\, U_4(x_1)$ must be a string in $\mathscr{L}_\varphi$. However, since $U_3(x_1)\, U_4(x_1) \neq U_3(x_2)\, U_4(x_2)$ and $d_1 + d_2 \geqslant 10kt$, $\hat{U}_1 U_2(x_2)\, U_3(x_1)\, U_4(x_1)$ is not of the form $xb\varphi(x)\, bx$. It is a contradiction.

Proof of Case 2. Unless stated otherwise, the notation used in the following are the same as we have used in the proof of the Case 1. In Case 2, for at least a quarter of all strings $Y(x)$'s in $\alpha_0{}^t$, the lengths of $Y_1^{j_0}(x)$ are less than t, or the lengths of $Y_2^{j_0}(x)$ are less than t. We shall show the contradiction in the former case (we can show similarly the contradiction in the latter case). Let $\alpha_1{}^t$ be a set of all strings of $\alpha_0{}^t$ such that the length of $Y_1^{j_0}(x)$ is less than t. Then $\#(\alpha_1{}^t) \geqslant 2^{19kt/2}$. For each element $Y(x)$ of $\alpha_1{}^t$, we define $Q(x) = Y_1^{j_0}(x)\, y_1^{j_0}(x)\, Y_c^{j_0}(x)$. Let $\alpha_2{}^t$ be the largest subset of $\alpha_1{}^t$ satisfying the condition such that for each $Y(x_1)$ and $Y(x_2)$ in $\alpha_2{}^t$, $Q(x_1) = Q(x_2)$. Then $\#(\alpha_2{}^t) \geqslant 2^{19kt/2}/(s+1)(2^k)^{2t+s+1} \geqslant 2^{14kt/2}$. Using considerations analogous to the case 1, for any $p_3 \gneqq 0$ and sufficiently large t, there exists a point e in the segment $Y_c^{j_0}(x)$ satisfying the inequality $\rho_e \lneqq p_3 t$. Let $\alpha_3{}^t$ be the largest subset of $\alpha_2{}^t$ whose elements satisfy the inequality $\rho_e \lneqq p_3 t$, where e is an appropriate fixed point of $Y_c^{j_0}(x)$.

Let $\alpha_4{}^t$ be the largest subset of $\alpha_3{}^t$ such that every element $Y(x)$ of $\alpha_4{}^t$ has the same trace of the derivation $D_x{}'$ at the point e. Then $\#(\alpha_4{}^t) \geqslant 2^{12kt/2}$. For an arbitrary string $Y(x)$ in $\alpha_4{}^t$, we represent $Y^{j_0}(x)$ in the form $Y^{j_0}(x) = z_1(x)\, z_2(x)$, where the interface point between $z_1(x)$ and $z_2(x)$ is e. Since $\alpha_4{}^t$ is a subset of $\alpha_2{}^t$, the segments $z_1(x)$'s are independent on x for all $Y(x)$ in $\alpha_4{}^t$. Therefore we may denote $z_1(x)$ by $\hat{z}_1$ and obtain $Y^{j_0}(x) = \hat{z}_1 z_2(x)$. The exact descendant of the segment $z_i(x)$ in $Y(x)$ is denoted by $U_i(x)(i = 1, 2)$. Let $\alpha_5{}^t$ be the largest subset of $\alpha_4{}^t$ such that for any $Y(x_1)$ and $Y(x_2)$, $|\, U_1(x_1)| = |\, U_1(x_2)|$. Then we have the inquality (16):

$$\#(\alpha_5{}^t) \geqslant 2^{11kt/2} \tag{16}$$

For all elements $Y(x)$'s in $\alpha_5{}^t$, $|\, U_1(x)| \geqslant 10kt$, or $|\, U_2(x)| \geqslant 10kt$. Without loss of generality we may assume that for all $Y(x)$'s, $|\, U_2(x)| \geqslant 10kt$ (the proof for the former case is similar). If every element $Y(x)$ in $\alpha_5{}^t$ has the same $U_1(x)$, then, $\#(\alpha_5{}^t) \leqslant 2^{5kt}$ because $|\, U_1(x)| \geqslant 5kt$. From (16), it is a contradiction. Therefore there must exist $Y(x_1)$ and $Y(x_2)$ in $\alpha_5{}^t$ such that $U_1(x_1) \neq U_1(x_2)$. Let $Y(x_1)$ and $Y(x_2)$ be strings in $\alpha_5{}^t$ such that $U_1(x_1) \neq U_1(x_2)$. Then we have (17) and (18):

$$\hat{z}_1 z_2(x_1) \overset{*}{\Rightarrow} U_1(x_1)\, U_2(x_1) = Y(x_1) \tag{17}$$

$$\hat{z}_1 z_2(x_2) \overset{*}{\Rightarrow} U_1(x_2)\, U_2(x_2) = Y(x_2). \tag{18}$$

Applying Lemma 4 at the point e in (17) and (18), we get $\hat{z}_1 z_2(x_1) \overset{*}{\Rightarrow} U_1(x_2)\, U_2(x_1)$. Hence $U_1(x_2)\, U_2(x_1)$ must be in $\mathscr{L}_\varphi$. But it is a contradiction.

From the above argument, there is no T-derivable CSEG G such that $L(G) = \mathscr{L}_\varphi$, where T satisfies the condition (1). Hence from Lemma 3, L_T does not contain $\mathscr{L}_\varphi$.

Q.E.D.

The following corollary is obtained immediately from Theorem 5.

COROLLARY 6. *Let $\hat{\varphi}$ be a single valued function defined on the set of all strings on an alphabet $\Sigma = \{a_1, a_2\}$, taking as values strings on the same alphabet Σ and satisfying the condition such that for all strings x in $\{a_1, a_2\}^*$, $|\, \hat{\varphi}(x)| \leqslant k\, |\, x\, |$, where k is a constant. Let $\hat{\mathscr{L}}_\phi = \{xb\hat{\varphi}(x)\, bx \mid x \in \{a_1, a_2\}^*\}$, where b is a symbol not in Σ. If T is a time function such that $\lim_{n\to\infty} (T(n)/n^2) = 0$, $\hat{\mathscr{L}}_\phi$ cannot be contained in L_T.*

A. V. Gladkij's Nonlinear Theorem on CSG's is the proposition obtained by replacing L_T by $L_{\langle\langle T\rangle\rangle}$ in Corollary 6. The open problem proposed by R. V. Book is the proposition obtained by replacing L_T by $L_{\langle T\rangle}$ in Corollary 6. Since $L_T \supseteq L_{\langle T\rangle} \supseteq L_{\langle\langle T\rangle\rangle}$, Corollary 6 is stronger than A. V. Gladkij's Nonlinear Theorem on CSG's and the open problem proposed by R. V. Book. Therefore our Corollary 6 solved R. V. Book's open problem.

Some closure and nonclosure properties on L_T, $L_{\langle T\rangle}$, and $L_{\langle\langle T\rangle\rangle}$ are given in [11], [2], [1] and [14]. We now consider other nonclosure properties on L_T, L_{-T-}, $L_{\langle T\rangle}$ and $L_{\langle\langle T\rangle\rangle}$.

The following two lemmas were already established.

LEMMA 7. *For any time function* T, (i) $L_T, L_{-T-}, L_{\langle T\rangle}$, *and* $L_{\langle\langle T\rangle\rangle}$ *are closed under intersection with regular sets,* (ii) $L_T, L_{-T-}, L_{\langle T\rangle}$, *and* $L_{\langle\langle T\rangle\rangle}$ *are closed under* ϵ*-free homomorphism and* (iii) $L_T, L_{-T-}, L_{\langle T\rangle}$, *and* $L_{\langle\langle T\rangle\rangle}$ *are closed under union.* [2], [14]

Although closure properties on L_{-T-} in the above lemma were not given in the references, the proofs are quite similar to those of the others.

LEMMA 8. *The class of context-free languages are properly contained in* $L_{\langle\langle n\rangle\rangle}$. [1]

For each $\alpha \subseteq \Sigma^*$ *and* $\beta \subseteq \Sigma^*$, $\mathrm{shuf}(\alpha, \beta) = \{w_1 y_1 w_2 y_2 \cdots w_n y_n \mid n \geqslant 1, w_1 \cdots w_n \in \alpha, y_1 \cdots y_n \in \beta\}$. *If* $x = a_1 \cdots a_n$ *and each* a_i *is in* Σ, $x^R = a_n \cdots a_1 \cdot x^R$ *is called the reverse of* x.

THEOREM 9. *Let* T *be a time function such that*

$$\operatorname*{Inf}_{n\to\infty} (T(n)/n) \gneq 0 \qquad \text{and} \qquad \lim_{n\to\infty} (T(n)/n^2) = 0.$$

Then

(i) $L_T, L_{-T-}, L_{\langle T\rangle}$, *and* $L_{\langle\langle t\rangle\rangle}$ *are not closed under intersection;*

(ii) $L_T, L_{-T-}, L_{\langle T\rangle}$, *and* $L_{\langle\langle T\rangle\rangle}$ *are not closed under complement;*

(iii) $L_T, L_{-T-}, L_{\langle T\rangle}$, *and* $L_{\langle\langle T\rangle\rangle}$ *are not closed under* Shuf.

Proof. We shall prove the theorem for L_T, the proofs of the other statements being similar. For $L_{\langle T\rangle}$ and $L_{\langle\langle T\rangle\rangle}$, both (i) and (ii) are established in [2].

(i) Let $\Sigma = \{a, b\}$.

$\alpha_1 = \{wcw^R \mid w \in \Sigma^*\}\, c\Sigma^*$ and $\alpha_2 = \Sigma^* c\{wcw^R \mid w \in \Sigma^*\}$ are context-free languages, where c is a simbol not in Σ. From Lemma 8, α_1 and α_2 are in $L_{\langle\langle n\rangle\rangle}$ and thus in L_T, where T is any time function satisfying the condition of the theorem. From Corollary 6, $\alpha_1 \cap \alpha_2 = \{wcw^R cw \mid w \in \Sigma^*\}$ is not contained in L_T. Therefore L_T is not closed under intersection.

(ii) Let $\bar{\alpha}$ be the compliment of α.

$\alpha_1 \cap \alpha_2 = \overline{(\bar{\alpha}_1 \cup \bar{\alpha}_2)}$. Therefore from (iii) of lemma 7 and (i) of this theorem, L_T is not closed under complement.

(iii) For $L_{\langle T\rangle}$ and $L_{\langle\langle T\rangle\rangle}$, (iii) follows (i) and the fact that these classes are AFL's (abstract families of languages) [2]. For L_T, the proof is as follows: Let $\Sigma = \{a, b\}$, and let $\hat{\Sigma} = \{\hat{a}, \hat{b}\}$. Let h be an ϵ-free homomorphism such that for each $x \in \Sigma$, $h(x) = x$ and for each $\hat{x} \in \hat{\Sigma}$, $h(\hat{x}) = x$. Then

$$\begin{aligned} &h(\mathrm{Shuf}(\beta_1 = \{wcw^R \mid w \in \Sigma^*\}, \beta_2 = \{wcw^R \mid w \in \Sigma^*\}) \cap \Sigma^* c\{x\hat{x} \mid x \in \Sigma\}^* c\hat{\Sigma}^*) \\ &\quad = \{wc\tilde{w}cw \mid w \in \Sigma^*\}, \end{aligned}$$

where if $w = x_1 \cdots x_n$ and each x_i is in Σ, $\tilde{w} = x_n{}^2 \cdots x_1{}^2$. From Corollary 6, $\{wc\tilde{w}cw \mid w \in \Sigma^*\}$ is not contained in L_T. Since β_1 and β_2 are context free languages, they are in L_n and thus in L_T. Therefore, from (i) of Lemma 7 and (ii) of Lemma 7, L_T is not closed under Shuf. Q.E.D.

Next we shall show some relations among several classes of languages.

DEFINITION 17. (i) A language α is called a quasirealtime language if and only if there exists a nondeterministic multi-tape Turing machine which accepts α in real time. [4] (ii) A language α is called a one-way nondeterministic stack language if and only if there exists a one-way nondeterministic stack automaton which accepts α. [9] (iii) A language α is called a nondeterministic realtime stack language if and only if there exists a one-way nondeterministic stack automaton which accepts α in real time. (iv) A language α is called a one-way deterministic stack language if and only if there exists a one-way deterministic stack automaton which accepts α. (v) A language α is called a deterministic realtime stack language if and only if there exists a one-way deterministic stack automaton which accepts α in real time. (vi) A language α is called a realtime language if and only if there exists a multi-tape Turing machine which accepts α in real time. (vii) A language α is called a k-tape realtime language if and only if there exists a Turing machine with k working tapes which accepts α in real time. (viii) A language α is called a deterministic realtime list-storage language if and only if there exists a one-way deterministic list-storage automaton which accepts α in real time. [16] (ix) A language α is called a nondeterministic realtime list-storage language if and only if there exists a one-way nondeterministic list-storage automaton which accepts α in real time. [10]

The following four lemmas are well known.

LEMMA 10. *There exists an ϵ-free* CFL *which is not a realtime language.* [17].

LEMMA 11. *The class of one-way nondeterministic stack languages is closed under homomorphisms.* [9]

LEMMA 12. *Each one-way nondeterministic stack language is a recursive set.* [9]

LEMMA 13. *For an arbitrary* PSL α, *there exist an n-bounded* CSL β *and a homomorphism h such that* $\alpha = h(\beta)$. [2]

The next lemma is clear.

LEMMA 14. $\{wcw^Rcw \mid w \in \Sigma^*\}$ *is a deterministic realtime stack language* (*and thus a* 1-*tape realtime language*) *and a deterministic realtime list-storage language, where c is a symbol not in* Σ.

THEOREM 15. *Let T be a time function such that* $\mathrm{Inf}_{n\to\infty}(T(n)/n) \gneq 0$ *and* $\lim_{n\to\infty}(T(n)/n^2) = 0$. *Then there is no inclusion relation between any class of* L_T, L_{-T-}, $L_{\langle T\rangle}$, *and* $L_{\langle\langle T\rangle\rangle}$, *and any class of the following classes*: (i) *The class of one-way nondeterministic stack languages.* (ii) *The class of nondeterministic realtime stack languages.* (iii) *The class of one-way deterministic stack languages.* (iv) *The class of realtime languages.*

Proof. This theorem follows from Corollary 6, Lemma 10, Lemma 11, Lemma 12, Lemma 13 and Lemma 14. Q.E.D.

Thus we have the next corollary immediately. Essentially this corollary is given in [2].

COROLLARY 16. *Let T be a time function such that* $\mathrm{Inf}_{n\to\infty}(T(n)/n) \gneq 0$ *and* $\lim_{n\to\infty}(T(n)/n^2) = 0$. *Then L_T does not include the class of nondeterministic realtime list-storage languages. Therefore L_T does not include the class of quasirealtime languages.*

It is known that the class of quasirealtime languages properly includes L_n [14]. A language α is quasirealtime if and only if there exist CFL's β, γ, and θ, and a length preserving homomorphism h such that $\alpha = h(\beta \cap \gamma \cap \theta)$ [4]. $L_{\langle\langle n^2\rangle\rangle}$ contains any CFL. is closed under length-preserving homomorphisms and is closed under intersection [2]. Hence $L_{\langle\langle n^2\rangle\rangle}$ includes the class of quasirealtime languages. Therefore Corollary 16 gives an interesting result that the class of quasirealtime language is included in $L_{\langle\langle n^2\rangle\rangle}$ but not included in L_T, where $\lim_{n\to\infty}(T(n)/n^2) = 0$.

4. CHAINS OF DISTINCT COMPLEXITY CLASSES

It is known that there exists an infinitely long chain, $L_T \subsetneq L_{T_2} \subsetneq \ldots$, of distinct complexity classes. In this section we shall show some infinitely long chains of distinct complexity classes between two complexity classes.

DEFINITION 18. Let

$$\beta_r = \{c^k xcx \mid x \in \Sigma^*, \lceil \log_2 k\rceil \geqslant r\lceil \log_2 |x|\rceil,$$

where $\Sigma = \{a_1, a_2\}$, c is a symbol not in Σ, r is a rational number in the range $1 \leqslant r \lneq 2$ and $\lceil s\rceil$ is the smallest integer m satisfying the condition $s \lneq m$.

LEMMA 17. *Let $T(n)$ be a time function such that* $\lim_{n\to\infty}(T(n)/n^{2/r}) = 0$. *Then β_r is not in L_T.*

Proof. Assume that β_r is in L_T. Then we will be lead to a contradiction. Let $G = (V, \Sigma \cup \{c\}, P, \sigma)$ be a $T(n)$-derivable grammar such that $\beta_r = L(G)$. Let $G' = (V \cup \hat{\Sigma} \cup \{\hat{c}, \#, \tilde{\#}, \sigma_0\}, \ \Sigma \cup \{c\}, \ \hat{P}, \sigma_0)$, where $\hat{\Sigma} = \{\hat{a} \mid a \in \Sigma\}, \#, \tilde{\#}, \sigma_0, \hat{c}$

and each $\hat{a}$ in $\hat{\Sigma}$ are new symbols not in V, respectively, and $\hat{P}$ is defined as follows: Let h be a homomorphism from V to $V \cup \hat{\Sigma} \cup \{\hat{c}\}$ such that for each a in $\Sigma \cup \{c\}$, $h(a) = \hat{a}$ and for each A in $V - (\Sigma \cup \{c\})$, $H(A) = A$. $\hat{P} = \{\xi \to h(\eta) \mid \xi \to \eta \in P\} \cup \{\sigma_0 \to \# h(\eta) \# \mid \sigma \to \eta \in P\} \cup \{\#\hat{a} \to a\tilde{\#} \mid a \in \Sigma\} \cup \{\tilde{\#}\hat{a} \to a\tilde{\#} \mid a \in \Sigma\} \cup \{\#\hat{c} \to \#, \tilde{\#}\hat{c} \to c\tilde{\#}, \tilde{\#}\# \to \epsilon, \#\hat{c}\# \to c\}$. Then $L(G') = \{wcw \mid w \in \Sigma^*\}$. From the definition of β_r, if $\lceil \log_2 k \rceil \geqslant r\lceil \log_2 |x| \rceil$, $c^k xcx$ is in β. Therefore, $c^i xcx$ is in β_r, where $i = [(2 \mid x \mid)^r]$. Since G is T-derivable,

$$\operatorname{dlm}(\sigma \overset{*}{\Rightarrow}_G c^i xcx) \leqslant T([(2 \mid x \mid)^r] + 2 \mid x \mid + 1) \leqslant T([(5 \mid x \mid)^r]).$$

Hence

$$\operatorname{dlm}(\sigma_0 \overset{*}{\Rightarrow}_{G'} xcx) \leqslant T([(5 \mid x \mid)^r]) + [(2 \mid x \mid)^r] + 2 \mid x \mid + 2.$$

Since $\lim_{n\to\infty} (T(n)/n^{2/r}) = 0$ and $r \lneq 2$, there exists a time function T' such that $\lim_{n\to\infty} (T'(n)/n^2) = 0$ and G' is T'-derivable. From Corollary 6, this is a contradiction. Q.E.D.

LEMMA 18. *Let $T_r(n) = n^{2/r}$, where r is a rational number in the range $1 \leqslant r \lneq 2$. Then β_r is in $L_{\langle\langle T \rangle\rangle}$.*

Proof. We can construct an $n^{2/r}$-bounded CSG G such that $L(G) = \beta_r$. Since the strictly formal proof is straightforward and quite tedious, we shall describe the outline of the proof. Let $r = p/q$, where p and q are positive integers, respectively. Let $\hat{a}_1$, $\hat{a}_2$ and $\hat{c}$ be nonterminal symbols corresponding to the terminal symbols a_1, a_2 and c, respectively, and let $\hat{x}$ be a nonterminal word corresponding to a terminal word x for each x in $\{a_1, a_2\}^* - \{\epsilon\}$. CSG G produces each word in β_r as follows:

First G derives $c^k xcw$. Let $\eta(t)$ be the length of the string at the tth step of the derivation. Let $f(n) \simeq g(n)$ mean that $0 < \lim_{n\to\infty} (f(n)/g(n)) \lneq \infty$. Then the length of the derivation is $\simeq\eta(t)$ at each step up to the step where $c^k xcw$ is derived, that is $t \simeq \eta(t)$. Next G derives the binary expression of k and $|x|$ by making a series of passes of a special marker across the block c^k and the block x. The method is quite similar to the Hennie's method of determining logarithms in [12]. Notice that $\lceil \log_2 k \rceil$ is equal to the length of the binary expression of k. On each pass G marks the block c^k and the block x with on the first symbols, the third symbols, the fifth symbols, etc. previously unmarked symbols, respectively. If the length of c^k (the length of $|x|$) is odd, the first bit of the binary expression of $k(|x|)$ is 1. Otherwise, it is 0. If the number of the symbols marked at the ith pass in the block c^k (the block x) is odd, the $(i + 1)$th bit of the binary expression of $k(|x|)$ is 1. Otherwise, it is 0. G writes every p bits of the binary expression of k and every q bits of the binary expression of $|x|$ on an appropriate position of the block c^k as one symbol. Then G decides whether $\lceil \log_2 k \rceil \geqslant \lceil \log_2 |x| \rceil$. If this condition is satisfied, the derivation moves to the next step. Otherwise, the derivation stops. The length of the derivation at each step up to the final step of this checking is at most $\simeq\eta(t) \log_2 \eta(t)$. That is, t is at most

$\simeq\eta(t) \log_2 \eta(t)$. Then G checks up whether $\hat{x}$ and $\hat{w}$ are the same. If $\hat{x}$ and $\hat{w}$ are the same, G changes $\hat{c}^k\hat{x}\hat{c}\hat{x}$ to c^kxcx. Otherwise, the derivation stops. The number of steps for this checking is at most $\simeq |x|^2$. Therefore, for each t, t is at most $\simeq\eta(t)^{2/r}$. Hence the above G is $n^{2/r}$-bounded CSG. That is, β_r is $n^{2/r}$-bounded CSL. Q.E.D.

The next two theorems are straightforward from Lemma 17 and Lemma 18.

THEOREM 19. *Let p and q be any real numbers such that $1 \leqslant p < q \leqslant 2$. Then* (i) $L_{n^p} \subsetneq L_{n^q}$, (ii) $L_{-n^p-} \subsetneq L_{-n^q-}$, (iii) $L_{\langle n^p\rangle} \subsetneq L_{\langle n^q\rangle}$, *and* (iv) $L_{\langle\langle n^p\rangle\rangle} \subsetneq L_{\langle\langle n^q\rangle\rangle}$.

THEOREM 20. *Let p and q be any real numbers such that $1 \leqslant p \lneqq q \leqslant 2$. Then there exist infinitely long chains of distinct complexity classes as follows*:

(i) $L_{n^p} \subsetneq L_{T_2} \subsetneq L_{T_3} \subsetneq \cdots \subsetneq L_{n^q}$,

(ii) $L_{-n^p-} \subsetneq L_{-T_2-} \subsetneq L_{-T_3-} \subsetneq \cdots \subsetneq L_{-n^q-}$,

(iii) $L_{\langle n^p\rangle} \subsetneq L_{\langle T_2\rangle} \subsetneq L_{\langle T_3\rangle} \subsetneq \cdots \subsetneq L_{\langle n^q\rangle}$,

(iv) $L_{\langle\langle n^p\rangle\rangle} \subsetneq L_{\langle\langle T_2\rangle\rangle} \subsetneq L_{\langle\langle T_3\rangle\rangle} \subsetneq \cdots \subsetneq L_{\langle\langle n^q\rangle\rangle}$.

From Theorem 19 and Theorem 20 the similar results on the hierarchies of time complexities of nondeterministic off-line two-pushdown-tape automata (abbreviated off-2pda's) are immediately derived. An off-2dpa is an acceptor which has finite internal states and two pushdown-tapes. An input data is initially given on one of two pushdown-tapes. We can consider that an off-2dpa is equivalent to a nondeterministic off-line list-storage tape automaton. [13] It is clear that if $T(n)$ is a time function satisfying the condition $\inf_{n\to\infty} (T(n)/n) \gneqq 1$, L_T coincides with the class of $T(n)$-recognizable languages by off-2pda's. Therefore we have the following two theorems.

THEOREM 21. *Let p and q be any real numbers such that $1 < p \lneqq q < 2$. Then the class of n^p-recognizable languages by off-2pda's is properly contained the class of n^q-recognizable languages by off-2pda's.*

THEOREM 22. *Let p and q be any real numbers such that $1 < p < q < 2$. Then there exists an infinitely long chain of distinct complexity classes as follows*:

$$\langle \mathscr{L}_{n^p} \rangle \subsetneq \langle \mathscr{L}_{T_2} \rangle \subsetneq \langle \mathscr{L}_{T_3} \rangle \subsetneq \ldots \subsetneq \langle \mathscr{L}_{n^q} \rangle,$$

where $\langle \mathscr{L}_T \rangle$ is a class of languages recognizable by off-2pda's of time complexity T.

Recently S. T. Cook has shown that for any real numbers $1 \leqslant r \lneqq s$, the class of languages recognizable by nondeterministic off-line multitape Turing machines running in $T_1(n) = n^r$ is properly contained in the corresponding class recognizable by $T_2(n) = n^s$ [5].[1] It is well known that the number of operations of a 1-tape Turing

[1] The authors were informed this by the referee.

machine needs to simulate $T(n)$ operations of a k-tape Turing machine is at most $T(n)^2$. The corresponding result on nondeterministic Turing machines is the same. It is clear that a derivation of length $T(n)$ in a PSG G is simulated by $cT(n)$ steps of a nondeterministic off-line multitape Turing machine, where c is a constant independent n. It is also well known that if M is a $T(n)$-time bounded nondeterministic off-line 1-tape Turing machine, then there is a $T(n)$-bounded PSG G such that $L(G) = L(M)$, where $T(n)$ is a time function and $L(M)$ is a set of languages accepted by M. [2] Combining these results, we can obtain the following results.

THEOREM 23. *Let p and q be any real numbers such that $1 \leqslant p \lneq q/2$. Then L_{n^p} is properly contained* $L_{\langle n^q \rangle}$.

COROLLARY 24. *Let p and q be any real numbers such that $1 \leqslant p \lneq q/2$. Then*

(i) L_{n^p} *is properly contained in* L_{n^q},

(ii) $L_{\langle n^p \rangle}$ *is properly contained in* $L_{\langle n^q \rangle}$.

ACKNOWLEDGMENT

It is pleasure to thank the referee of this paper for his many valuable suggestions.

REFERENCES

1. R. V. BOOK, Grammars with linear time functions, Proceeding of the Ninth Annual Symposium on Switching and Automata Theory, *Trans. IEEE*, New York, 1968, 298–305.
2. R. V. BOOK, Grammars with time functions, Harvard University Report, NSF-23, 1969.
3. R. V. BOOK, Time-bounded grammars and their languages, *J. Comput. System Sciences* **5** (1971), 397–429.
4. R. V. BOOK AND S. A. GREIBACH, Quasi-realtime languages, *Math. Systems Theory* **4** (1970), 97–111.
5. S. COOK, A hierarchy for nondeterministic time complexity, Fourth ACM Symposium on Theory of Computing, Denver, Colorado, 1972, 187–192.
6. M. DAVIS, "Computability and Unsolvability," McGraw-Hill, New York, 1958.
7. S. GINSBURG, "The Mathematical Theory of Context-Free Languages," McGraw-Hill, New York, 1966.
8. S. GINSBURG, S. A. GREIBACH, AND M. A. HARRISON, Stack automata and compiling, *J. Assoc. Comput. Mach.* **14** (1967), 172–201.
9. S. GINSBURG, S. A. GREIBACH, AND M. A. HARRISON, One-way stack automata, *J. Assoc. Comput. Mach.* **14** (1967), 389–418.
10. S. GINSBURG AND M. A. HARRISON, One-way nondeterministic real-time list-storage languages, *J. Assoc. Comput. Mach.* **15** (1968), 428–446.
11. A. V. GRADKIJ, On the complexity of derivations in immediate constituent grammars, *Algebra i Logika* **3**, **5** (1964), 29–44.

12. F. C. Hennie, One-tape off-line Turing machine computations, *Information and Control.* 8 (1965), 553–578.
13. Y. Igarashi and N. Honda, A classification of phrase structure languages, *Trans. Institute Electronics Communication Engineers Japan* 53-c (1970), 462–468.
14. Y. Igarashi and N. Honda, Some properties of $T(n)$-derivable languages, *Trans. Institute Electronics Communication Engineers Japan* 53-c (1970), 637–643.
15. Y. Igarashi, The complexity theory of formal languages, doctral dissertation, Tohoku University, Sendai, Japan, 1971.
16. Y. Igarashi and N. Honda, Deterministic real-time list-storage languages, *Trans. Institute Electronics Communication Engineers Japan* 53-c (1970), 514–521.
17. A. L. Rosenberg, Real-time definable languages, *J. Assoc. Comput. Mach.* 14 (1967), 645–662.